



Vers l'identification et le traitement des actes de dialogue composites

Frédéric Landragin

► To cite this version:

Frédéric Landragin. Vers l'identification et le traitement des actes de dialogue composites. JEP/TALN/RECITAL, Jun 2008, Avignon, France. pp.460-469, 2008. <halshs-00300235>

HAL Id: halshs-00300235

<https://halshs.archives-ouvertes.fr/halshs-00300235>

Submitted on 17 Jul 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Vers l'identification et le traitement des actes de dialogue composites

Frédéric Landragin

CNRS – Laboratoire LaTTICe (UMR 8094)
1 rue Maurice Arnoux – 92120 Montrouge
frederic.landragin@linguist.jussieu.fr

Résumé Il peut être difficile d'attribuer une seule valeur illocutoire à un énoncé dans un dialogue. En premier lieu, un énoncé peut comporter plusieurs segments de discours ayant chacun leur valeur illocutoire spécifique. De plus, un seul segment peut s'analyser en tant qu'acte de langage composite, regroupant par exemple la formulation d'une question et l'émission simultanée d'une information. Enfin, la structure du dialogue en termes d'échanges et de séquences peut être déterminante dans l'identification de l'acte, et peut également apporter une valeur illocutoire supplémentaire, comme celle de clore la séquence en cours. Dans le but de déterminer la réaction face à un tel acte de dialogue composite, nous présentons une approche théorique pour l'analyse des actes de dialogue en fonction du contexte de tâche et des connaissances des interlocuteurs. Nous illustrons sur un exemple nos choix de segmentation et d'identification des actes composites, et nous présentons les grandes lignes d'une stratégie pour déterminer la réaction qui semble être la plus pertinente.

Abstract Attributing one illocutionary value to a utterance in a dialogue can be difficult. First, a utterance can include several discourse segments, each one with a specific illocutionary value. Moreover, one discourse segment can be linked to a complex speech act that groups for instance a question together with the assertion of new information. Finally, the dialogue structure with the various exchanges and sequences can be decisive when identifying the dialogue act, and can also bring an additional illocutionary value, for instance the one consisting of closing the current sequence. With the aim to determine how to react to such a composite dialogue act, we present a theoretical approach to dialogue act analysis considering the task context and the dialogue participants' knowledge. We illustrate our choices in terms of segmentation and identification of composite acts, and we present the main features of a strategy for determining the most relevant reaction.

Mots-clés : Actes de langage complexes, structure du dialogue, terrain commun

Keywords: Complex speech acts, dialogue structure, common ground

1 Introduction

Le terme « acte de langage » désigne l'acte que l'on fait en parlant, en utilisant le langage dans une situation de communication. Dans la théorie de Searle, un acte de langage se définit par un but, un pré-requis, un corps (c'est-à-dire une réalisation), et des effets (Searle, 1969). Cette définition fait intervenir, en plus de la matérialisation linguistique, un certain nombre d'aspects relevant des motivations et des intentions communicatives. La communication, et par conséquent le dialogue et le dialogue homme-machine, font intervenir non seulement la compréhension des contenus des messages échangés, mais aussi la reconnaissance des intentions sous-jacentes et l'attribution d'états mentaux, de manière à interpréter correctement ces messages échangés. En fin de compte, on tend à inclure des aspects extra-linguistiques dans le processus d'interprétation et dans la notion d'acte de langage. Le contexte de communication devient un élément essentiel, et c'est peut-être pourquoi certains préfèrent parler d'« actes de parole », d'« actes de communication » ou encore d'« actes de dialogue ». L'acte se définit quasiment par son contexte de réalisation, l'acte de dialogue désignant l'unité minimale de communication dans un contexte dialogique.

En suivant cette voie et en l'adaptant au domaine du traitement automatique des langues, un « acte de dialogue homme-machine » se définit par l'unité minimale d'échange entre la machine et son utilisateur. C'est l'unité sur laquelle se base le gestionnaire de dialogue pour comprendre l'action de l'utilisateur et décider du comportement à adopter, de l'action à exécuter, de l'information à fournir et du message à envoyer au générateur de langue naturelle. Dans ce contexte, l'acte de dialogue homme-machine est une notion surtout opérationnelle. L'annotation de dialogues homme-machine en tenant compte de la progression de la tâche applicative constitue alors une méthode pertinente pour tester cette notion d'acte.

Qu'il se déroule entre deux humains ou entre un humain et une machine, un dialogue se caractérise par un certain nombre d'aspects qui peuvent être inclus dans la notion d'acte de dialogue, en la complexifiant d'autant, c'est-à-dire en la rendant « composite ». Il en est ainsi : des tours de parole ; des échanges que l'on peut organiser en échanges principaux et en sous-échanges ; des séquences particulières d'ouverture et de clôture ; et surtout de la finalité du dialogue qui dirige les interlocuteurs vers la co-construction d'un but commun, la résolution d'une tâche telle que trouver un billet de train pour aller d'un endroit à un autre, obtenir une information enregistrée dans une base de données, ou encore faire exécuter à la machine une série d'actions sur les objets qu'elle gère. Comme chaque message échangé peut être lié à plusieurs niveaux d'acte de dialogue (niveau du simple tour de parole, niveau de l'échange), on parlera d'actes multi-niveaux ou d'actes composites (Popescu-Belis, 2005).

La notion d'acte composite a déjà été bien étudiée (on pensera par exemple aux travaux de Levinson ou de Gazdar), en particulier pour rendre compte des actes de langage indirects, avec l'hypothèse qu'un énoncé tel que « *peux-tu me passer le sel ?* » ou « *avez-vous l'heure ?* » peut transmettre selon les interprétations et les contextes, soit une requête (acte purement indirect), soit à la fois une question et une requête (acte composite). La maxime originelle « quand dire, c'est faire » est ainsi reprise par (Kerbrat-Orecchioni, 2001) en « quand dire, c'est faire une chose sous les apparences d'une autre » (actes purement indirects) et « quand dire, c'est faire plusieurs choses à la fois » (actes composites). La notion d'acte composite reste cependant pour le moment très liée à celle d'acte indirect, et ne nous semble pas assez développée dans le cadre du dialogue. La majorité des travaux portant sur les actes de langage dans le dialogue s'intéressent à déterminer des typologies d'actes (Popescu-Belis, 2005), dont la mise en œuvre pose de nombreux problèmes sur les phénomènes observés et la cohérence des catégories (Traum, 2000). Un exemple de typologie

progressant constamment est celle de Harry Bunt et de ses étudiants, accessible à : <http://ls0143.uvt.nl/dit/>. Certains chercheurs incluent dans leur typologie des aspects liés à la structure du dialogue, avec par exemple la notion de paire adjacente telle que la paire question-réponse (Asher, Lascarides, 2003), mais l'obtention d'une typologie sans recouvrement ni manque reste très délicate. Une voie qui nous semble éclairante est la détermination des typologies de *types* d'actes, avec les éléments suivants : les actes partiels (Bunt, Schiffrin, 2007) ; les actes simples (Searle, 1969) ; les actes indirects (Searle, 1975) ; les actes composites (Bunt, 2000) ; les tentatives d'actes (Cohen, Levesque, 1995) ; ou encore, en revenant à Searle, les faux-actes (si le locuteur dit « *Es regnet* » avec non pas une intention de sens mais une intention d'exercice de prononciation allemande). On parlera de l'un ou l'autre de ces types d'actes selon que l'action de l'utilisateur est interrompue, aboutie, dérivée, multiple, tentée ou factice. Une autre voie consiste tout simplement à dépasser la notion d'acte de langage et à mettre en place d'autres théories pour rendre compte des phénomènes de dialogue. Par exemple, (Trognon, Brassac, 1988) montrent que la théorie des actes de langage n'est pas à même de fournir une théorie des enchaînements conversationnels et que les théories qui essayent de rendre compte des structures du dialogue avec des actes de langage n'y parviennent pas. Pour notre part, nous pensons que c'est possible, à condition de bien approfondir la notion d'acte composite.

A quel moment observe-t-on un acte de dialogue composite ? Tous les actes de dialogue sont-ils composites ? Jusqu'où peut-on aller dans l'identification des « sous-actes élémentaires » portés par un énoncé ? A quel moment doit-on s'arrêter dans cette identification ? Dans cet article, nous proposons quelques pistes pour répondre à ces questions, en posant et en étudiant les problèmes liés à la nature des actes composites (section 2), puis en proposant une méthode d'analyse et de traitement de ces actes par un gestionnaire de dialogue (section 3).

2 Nature des actes de dialogue composites

2.1 En partant de l'énoncé lui-même

Première possibilité : on définit un « acte de dialogue » par l'acte porté par un énoncé en tant que tour de parole dans son ensemble. Ainsi, tout ce que le locuteur dit sans être interrompu par l'interlocuteur est groupé dans un acte. Dans ce cas, un énoncé comportant deux phrases ou plus est forcément un « acte composite ». Par ailleurs, une phrase répartie sur deux énoncés (quand par exemple l'interlocuteur poursuit et complète ce que dit le locuteur) porte alors deux actes de dialogue qui rendent bien compte des étapes interactionnelles. Ceci entraîne la prise en compte d'actes partiels (Bunt, Schiffrin, 2007), nécessaires pour rendre compte de la prise en charge par le locuteur de la première partie de la phrase seulement. L'interlocuteur, quant à lui, prend à sa charge plutôt l'ensemble des deux parties, bien qu'il n'ait pas énoncé la première (à moins que sa complétion ne vienne en aide au locuteur et ne reflète pas sa propre intention communicative, ce qui conduit alors à un faux-acte). Si la complétion de l'interlocuteur ne correspond pas à l'intention préalable du locuteur, nous avons bien deux actes différents pour rendre compte de ce fait.

Deuxième possibilité : on définit un acte de dialogue par l'acte porté par une phrase. Dans ce cas il y a acte composite quand la phrase comporte plusieurs segments de discours. Il y a donc acte composite quand la phrase comporte un commentaire, une apposition, une clause subordonnée, une relative non restrictive, un épithète, et de manière générale tout phénomène intra-phrastique qui permet de décomposer la phrase en plusieurs intentions communicatives. C'est le cas en particulier des phénomènes étudiés dans un autre but par (Potts, 2003). Tous

ces cas sont également inclus dans la définition précédente. Par contre, en suivant ce principe, un énoncé comportant deux phrases s'analyse avec deux actes de dialogue distincts et non avec un acte composite. Les conséquences pour le gestionnaire de dialogue sont les mêmes en ce qui concerne la détermination des intentions de l'utilisateur, mais pas en ce qui concerne la détermination des réactions possibles de la machine. En effet, la gestion de deux actes de dialogue sans lien entre les deux ne permet pas au système autant de possibilités que la gestion d'un seul acte comportant plusieurs sous-actes ordonnés. Car, en effet, des critères linguistiques permettent de classer les sous-actes par ordre d'importance (1. principale, 2. subordonnée), ce qui permet un choix de réponse au seul acte premier ou à l'acte premier et à l'acte second. Par ailleurs, une phrase répartie sur plusieurs énoncés s'analyse avec un seul acte de dialogue, ce qui peut poser des problèmes puisque cela ne rend pas compte de manière simple de qui a dit quoi et de qui s'est impliqué dans quoi.

Troisième possibilité : on définit un acte de dialogue par l'acte porté par le segment de discours, comme le fait la SDRT (Asher, Lascarides, 2003). Dans ce cas, les phénomènes intra-phrastiques qui comportent plusieurs segments de discours ne constituent pas des actes composites, de même d'ailleurs que les phrases coordonnant deux questions, deux ordres ou deux assertions. Par contre, un segment simple comme « *avez-vous l'heure ?* » peut toujours être considéré comme un acte composite qui regroupe une question et une requête.

Compte tenu de l'importance que nous donnons aux enchaînements d'énoncés dans le dialogue, et compte tenu du fait qu'en dialogue homme-machine finalisé il est rare qu'un énoncé comporte plusieurs phrases, nous préférons la première solution. Par ailleurs, on peut ajouter aux aspects purement linguistiques les aspects paralinguistiques, en particulier la prosodie. Une phrase qui avec une prosodie conventionnelle constitue un acte simple, peut avec une prosodie inattendue constituer un acte composite. L'exemple typique est la phrase déclarative qui, avec une prosodie adéquate reste assertive, et, avec un contour interrogatif (montant ou simplement plat), peut devenir un acte composite groupant assertion et question.

2.2 En tenant compte du contexte situationnel immédiat

Une étape supplémentaire pour aborder le problème de la nature des actes de dialogue composites est de considérer la situation communicative immédiate, c'est-à-dire le contexte visuel à l'instant de l'énonciation, en incluant les objets visibles et les comportements du locuteur, de l'interlocuteur, ainsi que des éventuels intervenants extérieurs. Les paramètres dont il faut alors tenir compte sont les gestes co-verbaux, les regards, les postures, les attitudes, les expressions faciales, etc. Par exemple, un locuteur qui se gratte ou qui produit toute une série de gestes trahissant des hésitations et un manque de confiance peut, en même temps qu'il produit un acte de langage de type assertion, ajouter un acte de dialogue non verbal très significatif, plus proche de la question que de l'assertion. Cet exemple est similaire à celui concernant la prosodie, en passant du paralinguistique à l'extra-linguistique. Selon (Kerbrat-Orecchioni, 2001) c'est le mouvement de tête ou le geste d'ostension concomitant qui permet à l'énoncé « *peux-tu me passer le sel ?* » de valoir pour une requête.

Pour nous, cet énoncé linguistique à l'interrogative reste essentiellement une question, mais l'apparition d'un geste à l'instant de l'énonciation est fondamentale. Dans certains cas, le geste ne pourra être interprété que grâce à l'énoncé concomitant. Un geste purement déictique viendra ainsi permettre à l'interlocuteur de résoudre une référence à un objet ou à un lieu, sans pour autant avoir une influence sur la valeur illocutoire de l'énoncé dans son ensemble. Dans d'autres cas, on pourra attribuer très clairement au geste une valeur illocutoire : un geste peut fonctionner comme assertion (c'est le cas de la plupart des gestes quasi-linguistiques qui

émettent une information à destination de leur destinataire), comme question (une mimique interrogative, par exemple) ou comme ordre (le geste du professeur qui indique la porte à un élève, ou, dans le domaine des interfaces graphiques, le geste en forme de croix qui indique au système une action de suppression). Il nous semble donc légitime de parler d'« acte gestuel » au même niveau qu'« acte de parole », tous deux étant des « actes de dialogue ». Dans les exemples que nous venons d'évoquer, certains gestes se voient attribuer également un contenu propositionnel, comme le geste en forme de croix. D'autres comme la mimique interrogative apportent juste une valeur illocutoire à l'énoncé (on ne connaît pas l'objet de la question). Au final, l'interlocuteur doit confronter les différentes valeurs illocutoires. Ainsi, pour revenir au cas d'une question produite en même temps qu'un geste à valeur illocutoire d'ordre, l'interlocuteur se trouve face à deux valeurs illocutoires distinctes. Il peut alors :

- considérer qu'il y a incompatibilité (en particulier si les contenus propositionnels diffèrent, par exemple si la question ne porte pas sur le même objet que le geste directif) ;
- considérer que l'une des valeurs illocutoires prend le pas sur l'autre : dans notre cas, le geste, s'il ne vaut que pour sa valeur illocutoire d'ordre, impose cette valeur à l'énoncé dont la partie linguistique est alors considérée comme acte indirect ;
- considérer que les valeurs illocutoires sont complémentaires, autrement dit que l'énoncé dans sa globalité porte simultanément les deux valeurs, et qu'il s'agit donc d'un acte de dialogue composite.

L'interlocuteur doit ainsi faire preuve de capacités de fusion multimodale au niveau des valeurs illocutoires. Cette fusion peut être également dirigée par des aspects applicatifs précis.

2.3 En tenant compte de la structure de dialogue

Une troisième façon d'aborder le problème consiste à considérer l'énoncé courant comme un élément d'une structure dialogique. Plus que le moment de l'énonciation, c'est tout l'historique du dialogue dont il est ici question. Cet historique commence avec l'énoncé précédent, produit par l'interlocuteur, et relève donc de la paire adjacente, avec les paires question-réponse, assertion-accusé de réception, et bien entendu l'exemple classique de la phrase à l'impératif « *prends la rue Nationale* » qui constitue la réponse à la question « *comment puis-je aller au laboratoire ?* » et non un ordre. Autrement dit, la reconnaissance de l'acte de dialogue de l'énoncé courant ne tient pas compte uniquement de l'énoncé courant dans son contexte immédiat, mais tient compte également de la structure du dialogue. Si la seconde est essentielle, le premier ne doit pas pour autant être ignoré : il se peut toujours que l'interlocuteur réagisse à la forme qu'il peut juger trop directive.

La structure du dialogue peut être modélisée en parlant d'un dialogue régissant et de sous-dialogues incidents, d'enchâssements hiérarchiques de paires action-réaction, ou encore d'imbrications et de croisements non hiérarchiques mais plus libres d'actes élémentaires tels que des ordres, des questions et des assertions. C'est ainsi que (Portner, 2004) propose la gestion de piles de questions en cours de discussion (QUD, *questions under discussion*), d'ordres à exécuter (TDL, *to do list*) et d'informations assertées (CG, *common ground*, du moins dans son acception restreinte aux assertions). Le modèle hiérarchique genevois (Roulet et al., 1985) avec ses cinq niveaux peut donner des indications précieuses sur les phénomènes à l'origine de l'identification des sous-actes d'un acte composite. Le premier niveau est celui de l'acte de langage et a déjà été discuté dans la section 2.1. Le deuxième niveau est celui de l'intervention et s'intéresse à l'énoncé en tant qu'unité de tour de parole. C'est là qu'est

soulignée l'importance d'un acte directeur, qui donne à l'interaction sa valeur pragmatique dominante, par rapport aux actes subordonnés qui peuvent ne provoquer aucune réaction. Le troisième niveau est celui de l'échange, avec les exemples de paires adjacentes. Le modèle genevois met ici l'accent sur plusieurs phénomènes. En particulier, une évaluation de la réponse (« *bon* », « *ok* ») suit souvent la réponse à la question (donc trois interventions et pas seulement deux). Ensuite le fait que le découpage en échanges ne suit pas celui en interventions : on assiste souvent à « P_1 » (P pour proposition ou contenu propositionnel) suivi de « $P_2 ; P_3$ » suivi de « P_4 », où se trouvent un échange P_1 - P_2 et un échange P_3 - P_4 . Enfin, on peut considérer que tout énoncé est en soi un accusé de réception, dans la mesure où sa présence même prouve que le dialogue continue, c'est-à-dire fonctionne, même si c'est de manière incomplète. Cet argument va clairement dans le sens de la présence systématique d'actes composites dans le dialogue. Le quatrième niveau est celui de la séquence, qui se définit pour un même thème et inclut les phases d'ouverture, de développement et de clôture. Tout énoncé peut ainsi porter un acte de dialogue de type clôture de la séquence courante, ce qui ajoute une possibilité supplémentaire de sous-acte. (Bunt, 2000) analyse dans ce sens un simple « *merci* ». Plus formellement, une séquence se termine quand les piles QUD et TDL sont vides. Enfin, le cinquième niveau, celui de l'interaction, inclut les types de communication et les paramètres qui les définissent : continuité du groupe des participants, du cadre spatio-temporel, ainsi que des thèmes abordés. Ce niveau ne présente aucune variabilité en dialogue homme-machine finalisé et nous ne le retiendrons donc pas pour la détermination des composantes d'un acte composite.

2.4 En tenant compte de quelques aspects sociologiques

Une quatrième façon d'aborder le problème se fait en tenant compte de la nature des interlocuteurs et des enjeux sociologiques de leur dialogue, avec des aspects tels que attirer l'attention de l'interlocuteur sur quelque chose de particulier ; garder le contrôle du sujet et du déroulement du dialogue ; ne pas perdre la face ; garder une certaine supériorité hiérarchique par rapport à l'autre ; etc. (Merin, 1994) propose par exemple une méthode formelle pour la prise en compte de ce qu'il appelle des actes sociaux élémentaires. Dans la mesure où tout énoncé possède une part sociale élémentaire, on arrive à l'identification d'au moins un sous-acte social dans l'acte composite correspondant. Plus récemment, (Muller et al., 2006) incluent dans les aspects sociologiques les engagements et les obligations, et les intègrent dans un cadre de sémantique formelle, la SDRT (Asher, Lascarides, 2003).

3 Procédure de traitement des actes composites dans le dialogue

L'ensemble des aspects décrits précédemment n'est pas forcément applicable au dialogue homme-machine, certains nécessitant des outils de TAL qui ne sont pas encore réalisés. Ainsi, la détection automatique des segments de discours semble actuellement bien plus faisable que celle des phénomènes intra-phrastiques permettant de décomposer une phrase en plusieurs intentions communicatives. D'autre part, le dialogue homme-machine exploite très souvent un support visuel à travers une interface graphique qui apparaît à l'écran et qui permet un certain nombre de contrôles en plus ou à la place de ceux autorisés par la modalité vocale. Chacun de ces contrôles constitue un acte de dialogue potentiel, « acte de dialogue » étant ici employé dans une acception très large qui inclut les actions directes sur la machine, actions non verbales mais participant à l'interaction, avec des traces physiques analysables. Face à tous ces actes, un premier problème consiste à ordonner les différents sous-actes par ordre d'importance ou de pertinence, et un deuxième problème consiste à choisir le nombre n supérieur ou égal à 1 tel que le système décide de réagir aux n premiers actes. Nous allons

proposer quelques pistes pour un tel traitement automatique des actes de dialogue composites, en considérant un extrait de dialogue inspiré d'un scénario classique de réservation de billets de trains (U vaut pour l'utilisateur et S pour le système) :

U₁ : « *Je voudrais aller à Paris* »

S₁ : « *OK, voici les chemins pour aller à Paris* »

U₂ : « *Combien de temps par ce chemin qui est sans doute le plus court ?* »

S₂ : « *Une heure vingt, mais ce n'est pas le chemin le plus court* »

Nous nous intéressons à l'analyse de l'énoncé U₂, en la décomposant en cinq étapes.

3.1 Première étape : analyse littérale

Cette première étape regroupe l'analyse sémantique, la résolution en contexte des références, et l'attribution des valeurs illocutoires littérales. Elle permet d'obtenir une représentation sémantique et pragmatique initiale, sous la forme de structures de traits ou de formes logiques. Avec notre exemple, elle procède tout d'abord à l'analyse de la proposition principale de la phrase constituant U₂, avec la résolution de la référence multimodale « *ce chemin* » + geste (multimodalité complémentaire classique dans les systèmes multimodaux à forte composante langagière). En utilisant des pseudo formes logiques qui manquent de précision mais ont l'avantage de clarifier l'explication, on obtient ainsi la forme logique (1). La relative est pour sa part analysée en tant que commentaire et donc en tant qu'assertion (2). Quant au fait que l'énoncé en lui-même est un accusé de bonne réception du précédent, il permet d'obtenir (3) :

(1) DEMANDE (U, S, $\lambda x.x = \text{durée_trajet}(\text{trajet}_1)$)

(2) DIRE_QUE (U, S, $\text{trajet}_1 = \text{plus_court_chemin_vers}(\text{Paris})$)

(3) DIRE_QUE (U, S, $\text{compréhension_énoncé_précédent} = \text{OK}$)

Cette proposition (3) relève du niveau de l'échange dans le modèle hiérarchique genevois. Le fait de poser une question relative à des chemins prouve que l'utilisateur a correctement perçu la réaction précédente du système (que ce soit l'énoncé verbal ou l'affichage graphique des chemins, vu que l'énoncé U₂ comporte une action de référence à un chemin). Si le système avait énoncé « *OK, voici les chemins pour aller à Nancy* » ou s'il avait affiché des chemins n'aboutissant pas à Paris, il est clair que (3) et que l'acte de dialogue résultant auraient été très différents, au point de changer complètement le déroulement de l'interaction.

Pour rendre compte du niveau de la séquence, on devrait aussi ajouter un acte correspondant au fait que le thème courant, à savoir la satisfaction du désir de l'utilisateur de trouver un moyen d'aller sur Paris, n'est pas épuisé. En effet, le simple fait que U₂ soit une question montre que S₁ n'a pas suffi à l'utilisateur. Nous ignorerons ce niveau pour ne pas surcharger.

3.2 Deuxième étape : analyse des états intentionnels sous-jacents

La deuxième étape met en rapport la représentation obtenue avec les états intentionnels (désirs, croyances) de l'utilisateur, c'est-à-dire qu'elle active un certain nombre d'hypothèses en fonction du résultat de l'étape précédente. Nous utilisons un prédicat DESIRER pour les désirs et un prédicat CROIRE pour les connaissances, sachant qu'on peut considérer qu'une connaissance est une croyance avec la force la plus élevée. A la suite de (Steedman, Petrick,

2007), nous utilisons en fait pour les connaissances les trois prédicats : SAVOIR_QUE ; SAVOIR_SI et SAVOIR_VALEUR_DE :

- (4) DESIRER (U, SAVOIR_VALEUR_DE (U, durée_trajet (trajet₁)))
- (5) CROIRE (U, trajet₁ = plus_court_chemin_vers (Paris))
- (6) DESIRER (U, SAVOIR_SI (U, trajet₁ = plus_court_chemin_vers (Paris)))

L'hypothèse (4) se déduit directement de la question (1). Ensuite, on a soit l'hypothèse (5), soit l'hypothèse (6). En fonction de celle qui va être retenue, la réaction du système va varier. Ainsi, l'hypothèse (5) traduit « est-ce que l'utilisateur croit que c'est le chemin le plus court ? », ce qui correspond à un commentaire exprimant une croyance et requérant probablement une confirmation ou une infirmation. Pour sa part, l'hypothèse (6) traduit « est-ce que l'utilisateur veut savoir si c'est le chemin le plus court ? », ce qui correspond en quelque sorte à un commentaire indirect, en fait une autre question qui requiert une réponse. Autrement dit on aurait ici un acte indirect à l'intérieur d'un acte composite.

3.3 Troisième étape : confrontation aux connaissances du système

La troisième étape active les connaissances et les rôles du système qui peuvent intervenir compte tenu du résultat de la première étape. Elle constitue un pendant à la deuxième étape, mais du côté du système plutôt que de l'utilisateur. Dans notre cas :

- (7) SAVOIR_QUE (S, durée_trajet (trajet₁) = 80 min)
- (8) SAVOIR_QUE (S, trajet₁ = plus_court_chemin_vers (Paris))
- (8') SAVOIR_QUE (S, trajet₁ ≠ plus_court_chemin_vers (Paris))

La proposition (7) traduit le fait que le système connaît effectivement la réponse à la question posée (1), ce que l'on attend d'un tel système. Et les propositions (8) et (8') traduisent le fait que le système est effectivement capable de confirmer ou d'infirmer le commentaire (2). Dans notre exemple, le commentaire est faux et on active la proposition (8').

3.4 Quatrième étape : hiérarchisation des sous-actes de dialogue

La quatrième étape se fonde sur les précédentes pour déterminer les valeurs illocutoires profondes et ainsi l'ensemble des sous-actes constituant l'acte de dialogue de l'énoncé à analyser. Elle consiste également à classer les sous-actes identifiés. Si nous retenons l'hypothèse (5) et non (6), l'acte composite correspondant à U₂ se compose donc d'une question, de l'expression d'une croyance (qui se trouve vraie) et d'un accusé de bonne compréhension.

Chacun de ces actes a des conséquences sur la gestion du dialogue. La question entraîne une réponse, c'est-à-dire qu'en ajoutant un item à QUD, elle incite fortement le système à y réagir. L'expression d'une croyance qui est vraie n'a pas beaucoup d'incidence : si elle se trouve confirmée, une proposition sous forme de croyance devient une proposition sous forme de connaissance. Elle a en tout cas moins d'incidence qu'une croyance fausse qui pourrait être à l'origine d'une incompréhension dans le dialogue. Quant à l'accusé de bonne réception, il va dans le sens de la continuation du dialogue : si le système réagit en lien avec U₂, il sera implicitement traité (la situation et la pertinence des actes auraient été très différentes s'il n'y avait pas eu accusé de bonne réception).

Si nous nous focalisons sur la question et l'expression de la croyance, nous avons des possibilités d'évolution du dialogue et de la tâche très différentes. Compte tenu du fait que nous nous situons dans le cadre du dialogue homme-machine finalisé, il apparaît que répondre à la question est plus pertinent que confirmer la croyance. Ce point va être à l'origine des choix de la cinquième étape, et son importance doit être soulignée : nous avons ici considéré dans le contexte de dialogue que donner une valeur de réponse à une question est plus pertinent que confirmer une croyance, que cela correspond mieux aux rôles du système et à la tâche en cours. Cela ne veut pas dire pour autant que nous classons systématiquement des comportements tels que « fournir une réponse à une question », « confirmer une croyance » ou encore « infirmer une croyance » dans un ordre prédéfini. Au contraire, c'est lors de cette étape que la finalité du dialogue et les contraintes de la tâche peuvent s'exprimer. Chaque acte de dialogue peut aller dans le sens de la satisfaction de la tâche ou non : dans certains cas comme le nôtre, confirmer une croyance peut faire moins avancer la tâche que fournir une réponse à une question ; dans d'autres cas, par exemple si la croyance est fondamentale pour la tâche en cours, ce sera l'inverse.

3.5 Cinquième étape : détermination des sous-actes auxquels réagir

Enfin, la cinquième étape se charge de la détermination de la réponse du système, c'est-à-dire du choix du nombre de sous-actes auxquels réagir, ainsi que du choix des types de réactions (contenus des messages à générer, ou « quoi dire » et « quoi faire »). Elle consiste pour le système à choisir de répondre seulement à la question qui a été identifiée comme acte premier, ou à la question et à l'expression de la croyance. Si l'écart de pertinence entre l'acte premier et l'acte second est significatif, on ne retient que l'acte premier et on génère quelque chose comme DIRE_QUE (S, U, durée_trajet (trajet₁) = 80 minutes). Si, comme ici, on retient les deux actes, on génère S₂. Nous ne nous étendons pas sur les critères permettant ce choix. Ces critères peuvent en effet être très nombreux et font intervenir beaucoup d'aspects de la tâche applicative. Par exemple, si le système a pour priorité de vendre le billet le plus cher, il aura peut-être intérêt à ne pas infirmer une croyance de manière à laisser l'utilisateur dans son erreur. Autrement dit, il sera amené dans un tel cas à mentir par omission. Cette conclusion illustre en quoi la gestion des actes composites est importante et doit être réalisée avec attention.

4 Conclusion et perspectives

Nous avons souligné que les actes de langage dans le dialogue sont plus complexes qu'il n'y paraît, et que la notion d'acte composite semble essentielle et omniprésente, alors qu'elle n'est pas tant traitée dans la littérature. Notre principal objectif était ainsi de mettre l'accent sur des phénomènes auxquels le TAL a trop tourné le dos et qui pourraient pourtant être à la base de systèmes communicants performants. Nous avons ainsi proposé un ensemble de pistes pour un modèle d'identification et de traitement automatique des actes composites dans des dialogues finalisés. L'intérêt principal de ce travail est de montrer comment une prise en compte fine des actes de dialogue permet de concevoir un gestionnaire de dialogue capable de mettre en rapport les intentions de l'utilisateur, les buts du dialogue et ses propres contraintes de fonctionnement. Il montre à quel point l'identification des actes de dialogue et le raisonnement sont des processus très liés. Il montre également que les aspects coopératifs, y compris peut-être les fameuses maximes de Grice, relèvent pour une grande part de ces processus. Pour le moment, la principale perspective à ce travail est l'élaboration d'un modèle formel en exploitant un cadre existant tel que la SDRT. Une autre perspective est

expérimentale et consiste à faire évaluer par des sujets les réactions et réponses d'un système (ou d'un système simulé) face à des énoncés comportant des actes composites.

Références

- ASHER N., LASCARIDES A. (2003). *Logics of Conversation*. Cambridge : Cambridge University Press.
- BUNT H. (2000). Non-Problems and Social Obligations in Human-Computer Conversation. Actes de *Third International Workshop on Human-Computer Conversation*, Bellagio, Italy.
- BUNT H., SCHIFFRIN A. (2007). Defining Interoperable Concepts for Dialogue Act Annotation. Actes de *Seventh International Workshop on Computational Semantics*, Tilburg, The Netherlands, 16-27.
- CAELEN J., XUEREB A. (2007). *Interaction et pragmatique*. Paris : Hermès-Lavoisier.
- COHEN P.R., LEVESQUE H.J. (1995). Communicative Actions for Artificial Agents. Actes de *First International Conference on Multi-Agent Systems*, Cambridge : MIT Press, 65-72.
- KERBRAT-ORECCHIONI C. (2001). *Les actes de langage dans le discours*, Nathan Université.
- MERIN A. (1994). Algebra of Elementary Social Acts. Dans : TSOHATZIDIS S.L. (Ed.) *Foundations of Speech Act Theory*. London : Routledge, 234-263.
- MULLER P., MAUDET N., PREVOT L. (2006). Social Constraints on Rhetorical Relations in Dialogue. Actes de *Workshop on Constraints in Discourse*, Maynooth, Ireland, 2006.
- POPESCU-BELIS A. (2005). Dialogue Acts: One or More Dimensions? ISSCO Working Paper n° 62, University of Geneva.
- PORTNER P. (2004). The Semantics of Imperatives within a Theory of Clause Types. Dans : WATANABE K., YOUNG R.B. (Eds.) *Proceedings of Semantics and Linguistic Theory 14*. CLC.
- POTTS C. (2003). The Logic of Conventional Implicatures. PhD, University of California.
- ROULET E., AUCLIN A., MOESCHLER J., RUBATTEL C., SCHELLING M. (1985). *L'articulation du discours en français contemporain*. Berne : Peter Lang.
- SEARLE J. (1969). *Speech Acts*. Cambridge : Cambridge University Press.
- SEARLE J. (1975). Indirect Speech Acts. Dans COLE P., MORGAN J. (Eds.) *Syntax and Semantics, Vol. 3: Speech Acts*. Academic Press.
- STEEDMAN M., PETRICK R. (2007). Planning Dialogue Actions. *Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue*. Antwerp, Belgium, 265-272.
- TRAUM D. (2000). 20 Questions on Dialog Act Taxonomies. *Journal of Semantics* 17 (1), 7-30.
- TROGNON A., BRASSAC C. (1988). Actes de langage et conversation. *Intellectica* 2 (6), 211-232.